

Tågtrafikplanering med successiv tilldelning

Malin Forsgren, Martin Aronsson, Per Kreuger

Swedish Institute of Computer Science (SICS)

Box 1263, SE-164 29 Kista, Sweden

email: {malin,martin,piak}@sics.se

Mars 2009

SICS Technical Report T2009:04

ISRN: SICS-T-2004/04-SE

ISSN: 1100-3154

Sammanfattning

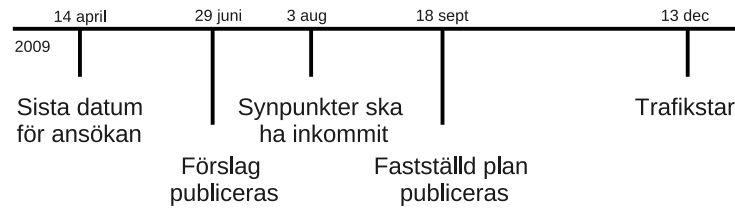
Idag konfliktregleras tågtrafiken i Sverige på sekundnivå redan i tidtabellägningsfasen då operatörernas ansökningar har kommit in till Banverket, trots hög osäkerhet i planeringsunderlaget. Den här rapporten beskriver en tidtabellägningsprocess där den detaljerade planeringen skjuts upp tills mer tillförlitliga fakta finns att tillgå. För den tidiga fasen föreslås i stället en mer översiktlig – grov – planering, vilken huvudsakligen mynnar ut i ett antal ankomster och avgångar som Banverket åtar sig att leverera punktligt till operatörerna. Poängen är att Banverket på så vis kan skjuta på besluten angående det som ska hända mellan åtagandepunkterna, och vi presenterar en konkret metod för hur detaljeringen som följer på den grova planeringen kan gå till. Metoden producerar en plan med en detaljeringsnivå som i stort överensstämmer med den i Banverkets tidtabelläggningssystem TrainPlan.

Keywords: Schemaläggning, tidtabelläggning, dynamisk planering

Detta arbete har finansierats genom Banverkets forskningsanslag HK 06-1522/AL50

Innehåll

1	Inledning	3
2	Principen bakom successiv tilldelning	4
2.1	Ett nytt sätt att planera	4
2.1.1	Ansökningarnas innehåll	4
2.1.2	Den grova planen	5
2.1.3	Åtagandet	7
2.1.4	Produktionsplanerna	8
2.1.5	Det slutgiltiga åtagandet	10
3	Demonstratorn	11
3.1	TP-databasen	12
3.2	Uppbyggnad av trafikrepresentationen	13
3.3	Domänbestämning	14
3.3.1	Indata	15
3.3.2	Beräkning av tider	16
3.4	Slackifiering	17
3.4.1	Generella principer	18
3.4.2	Beräkning	19
3.4.3	Allmän utvidgning	19
3.4.4	Algoritm	19
3.4.5	Specifik inskränkning	20
3.5	Filskrivning	21
3.5.1	Terminologi och beteckningar	21
3.5.2	Tågekvationer	21
3.5.3	Länkekvationer	22
3.5.4	Platsekvationer	23
3.5.5	Styv tidtabell och liknande krav	24
3.5.6	Domäner	24
3.5.7	Binärdeklaration	25
3.5.8	Objektfunktion	25
3.6	Optimering	25
4	Operatörernas ansökningar	27
5	Testfall	29
5.1	Åtagandet	30
5.1.1	Körning 1 – Åtagandet	30
5.2	Produktionsplaneringen	31
5.2.1	Körning 2 – Produktionsplanering	31
5.2.2	Om presentationen av resultaten	32
6	Sammanfattning	33
	Referenser	35



Figur 1: Tidplan för tilldelning av kapacitet gällande Tågplan 2010

1 Inledning

Tiden för att ansöka om att få köra tåg i Sverige år 2010 går ut i april 2009. Hela tidplanen för tilldelning av kapacitet gällande Tågplan 2010 framgår av Figur 1 [4]. När ansökningstiden går ut inleds ett arbete där Banverket under kort tid måste hitta en lösning som är godtagbar för alla parter. Ett viktigt led i arbetet är att bedöma om den tänkta trafiken är genomförbar eller inte. Metoden för det sistnämnda har hittills varit att hitta en konfliktfri tidtabell trots att det egentligen inte finns underlag för att planera på sekundnivå så långt i förväg pga att många uppgifter fortfarande är osäkra.

Det grundläggande problemet med dagens process är att man tidigt låser sig vid *en* lösning i stället för att hålla dörren öppen för flera alternativ. Då lösningen baseras på information vars detaljer i stor grad är okända för såväl operatörer som för Banverket vid planeringstillfället, kommer man göra både medvetna och omedvetna antaganden angående tågens prestanda för att kunna konfliktreglera i detalj. Konsekvensen blir att den plan som ska bevisa för alla inblandade att den föreslagna trafiken är genomförbar blott blir ett förslag till tidtabell bland många, för en trafik som endast *liknar* den trafik man egentligen behöver planera för. Planen inger alltså i själva verket en falsk trygghet.

Tidtabellläggning på Banverket sker helt utan något beräkningsverktyg som föreslår hur konflikter ska lösas och tar därför väldigt lång tid. I takt med att såväl beräkningsmetoder och -heuristiker som datorernas processorer har blivit kraftfullare, finns det idag större möjlighet att lösa stora schemalägningsproblem på kortare tid än för bara några år sedan. Om vi utnyttjar detta faktum ger det oss valmöjligheten att vänta längre med att börja planera i detalj, och därmed även ett incitament att se över hela processen. Huvudsyftet med den tidiga planeringsfasen är det som nämndes i början av detta kapitel: att ta fram en genomförbar plan som kan accepteras av alla inblandade. Vi ifrågasätter inte syftet, men ställer oss frågan om syftet kan uppnås på ett annat sätt än det som används idag.

Rapporten är uppbyggd på följande vis: vi inleder med en övergripande

beskrivning av det planeringskoncept som vi väljer att kalla successiv tilldelning. Därefter följer en omfattande beskrivning som på en konkret nivå steg för steg beskriver hur processen ser ut i vår demonstrator. Vi redogör sedan för problematiken kring att använda operatörernas ansökningshandlingar som utgångspunkt för optimering. Avslutningsvis redovisar vi våra erfarenheter av de inledande testkörningarna av den demonstrator som vi har tagit fram som en del av konceptet, samt avrundar med en sammanfattning av vad planering med successiv tilldelning innebär i praktiken.

2 Principen bakom successiv tilldelning

I kapitel 3 beskrivs ett konkret sätt att ta fram en optimerad plan från ett råmaterial. Den beskrivna processen har en central plats i ett förslag till nytt sätt att planera och tidtabellägga på. Det här kapitlet beskriver hur vi tänker oss det nya arbetssättet, där planering sker med *successiv tilldelning*. Vi använder ett grundligt men allmänt hållet exempel för att belysa de viktigaste aspekterna och några av fördelarna med det nya sättet jämfört med det gamla.

2.1 Ett nytt sätt att planera

Tågo, en fiktiv operatör på en avreglerad marknad, ansöker år 2015 precis som övriga operatörer om att få köra tåg på Banverkets spårnät. Tågo är intresserade av att så snabbt som möjligt få veta från Banverket om de får genomföra sin trafik enligt sina önskemål, alternativt på vilket sätt de måste kompromissa för att deras trafik ska fungera ihop med de andra operatörernas tåg. Ju snabbare och klarare besked Banverket kan ge, desto bättre blir Tågos förutsättningar för den fortsatta interna verksamhetsplaneringen.

2.1.1 Ansökningarnas innehåll

Vid ansökningstillfället har Tågo, som kör persontåg i en befolkningstät, medelstor region i Sverige (t ex Mälardalen), en idé om hur de vill köra tåg, men inte nödvändigtvis någon bestämd åsikt om exakt när deras tåg ska passera varennda namngiven plats i nätet. De planeringsansvariga på Tågo vet hur lång tid tågresorna maximalt bör få ta, hur ofta tågen ska gå, på vilka platser tågen ska släppa på och av passagerare, samt vid vilka platser många passagerare kommer att byta tåg eller byta till andra transportslag såsom bussar och färjor. Vissa tåg är dessutom tänkta att ingå i en styv tidtabell.

Bortsett från högprioriterade, specifika önskemål om avgångs- och ankomsttider som Tågo menar har stor betydelse för hur attraktiva deras tåg kommer att vara för kunderna, har Tågo inga synpunkter på hur tidtabellen kommer att se ut i detalj. Specifika önskemål av den typ vi syftar på ovan kan vara att ett tåg ska vara framme vid en viss station ett lämpligt antal minuter innan färjan avgår, eller avgå och ankomma vid vissa stationer på klockslag som passar pendlande resenärer. Anslutningar ska naturligtvis också respekteras i

möjligheste mån, vilket innebär att om ett visst tåg måste flyttas i tiden för att få rum, kanske andra tåg också måste flyttas för att viktiga anslutningar ska upprätthållas.

I det nya ansökningsförfarandet betonas att operatörerna bara kan få det de frågar om. Vidare uppmuntras operatörerna att å andra sidan endast ansöka om det som verkligen är viktigt, dvs inte vara mer specifika än deras trafikupplägg kräver eftersom Banverket behöver ett visst planeringsutrymme för att kunna möta samtliga operatörers önskemål på bästa sätt. I och med att operatörerna uppmanas rangordna sina önskemål, kan Banverket sikta på att tillgodose de högst prioriterade önskemålen, samtidigt som en operatör inte kan förvänta sig att de lägre prioriterade önskemålen blir uppfyllda om de står i konflikt med någon annans operatörs högt prioriterade planer.

I Tågos ansökning framkommer alla relevanta krav och önskemål, indelade i olika prioritetsskyltar. Tågo specificerar explicit i ansökningen alla anslutningar de har tänkt sig. Deras ansökan går innehållsmässigt att dela in i fyra kategorier, utöver önskemålet att överhuvudtaget få köra den mängd tåg som ingår i ansökningshandlingarna. Ansökan specificerar:

- Gångtider
- Avgångs- och ankomsttider
- Anknytningar
- Inbördes relationer (styv tidtabell)

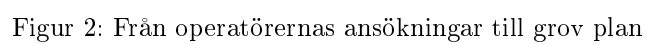
Gångtidsspecifikationerna säger hur lång tid resorna med deras olika tåg helst ska ta, med en angiven gräns för hur lång tid de maximalt får ta. Avgångs- och ankomsttider anges för relevanta platser, i prioritetsordning, som önskade klockslag tillsammans med intervall som anger vilka tidpunkter som är acceptabla om de önskade klockslagen inte går att förverkliga. Önskemålen om anknytningar och styv tidtabell anges som associationer mellan relevanta ankomster och avgångar tillsammans med ramar för hur förhållandena mellan dessa ska se ut för att önskemålen ska anses vara uppfyllda.

2.1.2 Den grova planen

På Banverket tar man hand om Tågos och alla andra operatörers ansökningar och börjar direkt arbeta med att hitta en fungerande plan som tar allas önskemål i beaktan på ett rättvist sätt. Målsättningen är att snabbt kunna ta fram en genomförbar, grov (översiktlig) plan för trafiken utan att behöva genomföra en planering som schemalägger minsta tågrörelse. Samtidigt behöver operatörerna få veta vilka av deras önskemål som kommer att bli uppfyllda och vilka som behöver kompromissas.

Med hjälp av den nya processen, för vilken Banverket har effektivt datorstöd, kan en grov plan arbetas fram. Se Figur 2.

Grundförutsättningarna för trafiken i form av realistiska gångtider för aktuella tågtyper och infrastrukturdata för aktuellt område och detaljeringsnivå



hämtas från en databas (motsvaras av TP-databasen i vår demonstratorbeskrivning i avsnitt 3.1). En datastruktur för trafiken och infrastrukturen som beräkningsverktyget kan hantera byggs upp (se avsnitt 3.2).

När ramarna för avgångs- och ankomsttiderna ska sättas, är utgångspunkten naturligtvis ansökningarna. Vissa platser i spårnätet är för vissa tåg associerade med mer eller mindre specifika klockslag medan andra helt saknar tidsangivelser. I det som motsvarar domänbestämningssteget (se 3.3) konkretiseras tiderna för alla händelser på de platser som är relevanta, *giltiga*, för planeringen så att varje sådan avgång och ankomst förknippas med ett intervall för när den kommer att kunna ske som är förenligt med de önskemål som tågets operatör faktiskt har angett i ansökan.

Tidtabellkonstruktörerna börjar med att undersöka om det finns en lösning som respekterar allas önskemål. Detta kan göras genom att optimera direkt på resultatet från föregående steg. Kriterierna för när den grova planeringen anses uppfylla alla krav och hur en sådan planering går till i detalj är inte fullständigt utrett ännu och beskrivs inte i denna rapport. Tillsvidare räcker det med att konstatera att det är fullt möjligt, om än osannolikt, att alla operatörers önskemål anses uppfyllbara utan minsta modifikation. Men om så är fallet är den grova planen givetvis klar.

Det troligaste är att operatörernas önskemål är i konflikt med varandra i åtminstone något avseende. Med hjälp av beräkningsverktyget kan tidtabellkonstruktörerna se vilka konflikterna är och få hjälp med att komma fram till vad som behöver göras åt dem, t ex genom att i omgångar tillämpa olika slack (se 3.4) och ta fram nya planer med färre och/eller mindre svårlösta konflikter.

Parallellt pågår en diskussion mellan tidtabellkonstruktörerna och operatörerna om vilka förändringar av specifikationen som är godtagbara. Det kan röra sig om att tillåta förändringar av enstaka tider eller flytta vissa tåg, förlänga tågs gångtider, ignorera vissa krav på anknytningar, helt stryka tåg ur trafikupplägget, etc.

Dialogen med operatörerna varvas med optimering av de reviderade förutsättningarna för trafiken tills godtagbara planer som dessutom fungerar ihop har hittats för alla tidtabellkonstruktionsområden.

2.1.3 Åtagandet

Efter att ha förankrat de största avvikelserna jämfört med ansökningarna hos berörda operatörer en sista gång, presenterar Banverket därefter delar av den grova planen som ett *åtagande* gentemot operatörerna. Åtagandet utgörs av vad Banverket lovar att leverera till operatörerna med en specificerad, hög sannolikhet.

Under processen då den grova planen arbetades fram, kom Tågo i samråd med tidtabellkonstruktörerna fram till att de förmodligen skulle bli tvungna att göra avkall på kravet på styv tidtabell för en grupp tåg på några av de mindre stationerna. Andra förändringar jämfört med ansökan blev ett antal flyttade tåg, några förlängda gångtider och några anslutningar som verkar bli lidande jämfört med det upplägg som Tågo hade tänkt sig. I övrigt är beskedet från

Banverket när åtagandet presenteras att Tågos trafik i stora drag kommer att kunna genomföras som planerat.

Banverkets åtagande i det här skedet, långt före operativ drift, talar inte i detalj om för operatörerna *hur* tågen ska ledas fram. Den talar däremot om för operatörerna *att* tågen, med hög sannolikhet, kommer att kunna befinna sig vid vissa stationer vid vissa, bestämda tidpunkter. Dessa besked ska vara tillräckligt säkra för att operatörerna ska kunna planera sina verksamheter kring dem och kunna publicera en tidtabell.

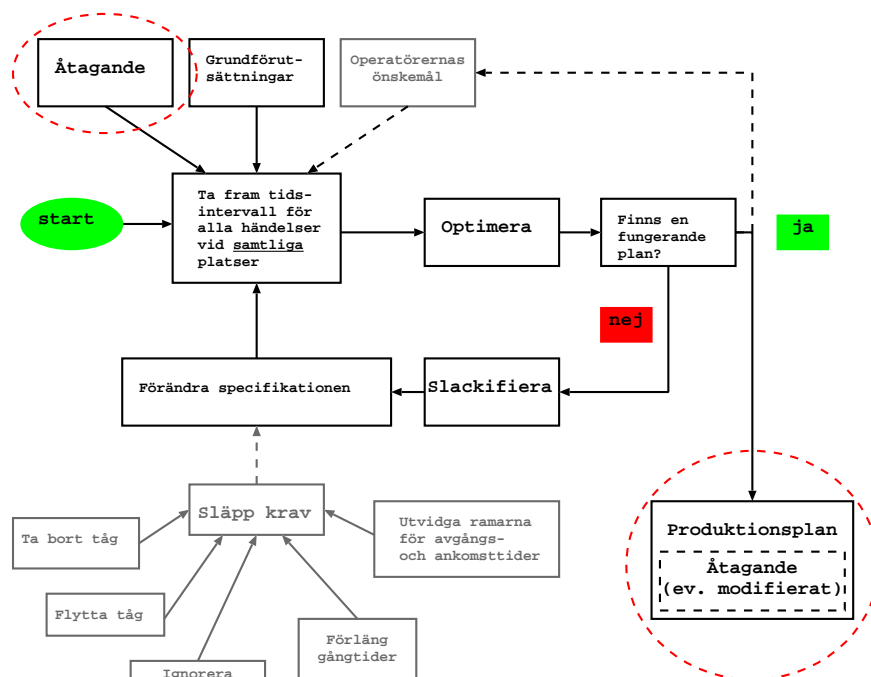
2.1.4 Produktionsplanerna

Dialogen mellan operatörerna och Banverket fortsätter även efter det att Banverket har offentliggjort sitt åtagande gentemot operatörerna. Den grova planen uppdateras med jämna mellanrum för att hållas ajour med det man för tillfället känner till om förutsättningarna för trafiken. Eftersom man har låst ankomster och avgångar endast vid en delmängd av punkterna i nätet, har man betydligt större möjligheter än förut att anpassa planen efter nya förhållanden – utan att göra avkall på de löften man har givit.

I vissa fall måste operatören vars uppgifter har förändrats finna sig i att planeringen inte kan använda de uppdaterade uppgifterna på ett konstruktivt sätt. Ett exempel är när ett tåg visar sig bli betydligt lättare än man först antog, i vilket fall tåget kan komma att behöva köra så långsamt som det ursprungligen planerades för om inte planen kan anpassas för att leda fram tåget snabbare utan att åtagandet för någon operatör bryts. Det finns även en risk att de tillkommer nya uppgifterna som är direkt oförenliga med de uppgifter som åtagandet baserades på, i vilket fall det kan bli nödvändigt att modifiera åtagandet. Hur det ska gå till är i så fall en förhandlingsfråga mellan Banverket och berörda operatörer.

Någon gång nära tidpunkten då den nya, detaljerade, dagliga tidtabellen ska tas i drift, begär Banverket slutgiltigt in de uppgifter från Tågo och de övriga operatörerna som ska ligga till grund för den detaljerade produktionsplanen. Operatörerna kan som nämnt inte räkna med att få nytillkomna önskemål tillgodosedda om de står i konflikt med det Banverket redan åtagit sig, men det är viktigt att inse att nya önskemål och förändrade förutsättningar inte nödvändigtvis skapar problem, utan även kan öppna för möjligheten att leda fram tågen mer effektivt. Ett uppenbart exempel på det sistnämnda är när operatörerna inte tänker köra alla tåg de har ansökt om, vilket givetvis lämnar större utrymme till den övriga trafiken.

Tågo kan så här nära den operativa driften lämna detaljerade och tillförlitliga uppgifter som tar bort mycket av osäkerheten i beräkningarna. Mindre osäkerhet i systemet leder till att man kan minska bufferterna eftersom man med större precision kan avgöra vad i planen som faktiskt är buffert (slack) och vad i planen som är nödvändigt för att kunna genomföra tågrörelserna. Som ett praktiskt exempel kan nämnas att sannolikheten att verkligen lyckas ringa in var ett tågmöte kommer att ske ökar om alla faktiska omständigheter kring tågen och infrastrukturen är kända. När man väl fastställt var ett tågmöte kommer att



Figur 3: Från grov plan till produktionsplan

äga rum, behöver man heller inte ta lika stor höjd för att det kan komma att ske någon annanstans.

Processen som leder fram till produktionsplanen sammanfattas i Figur 3 och har stora likheter med den process som tidigare har beskrivits och som mynnar ut i den grova planen och åtagandet (se Figur 2). Största skillnaderna är detaljeringsnivå respektive utgångspunkt. Eftersom produktionsplanen ska inbegripa tider på avgångar och ankomster vid samtliga platser (utom möjligtvis vid vissa blocksignaler och krysstationer), måste platserna vara explicit representerade när produktionsplanen ska tas fram. Även om åtagandet naturligtvis baseras på den grova planen som i sin tur baseras på de ursprungliga ansökningarna från operatörerna, är det åtagandet som står i fokus för produktionsplaneringen. Det i ansökningarna som operatörerna fortfarande vill ha, men som av något skäl inte ingår i åtagandet, ska däremot fungera vägledande i processen då produktionsplanen tas fram.

Bortsett från skillnaderna som just nämnts, liknar själva processen den process där den grova planen tas fram. Eftersom den grova planen bedömdes som genomförbar, undersöks först om den detaljerade planen helt ryms inom ramarna för åtagandet. Om så är fallet, kan önskemål från operatörerna som inte täcks in av det ursprungliga åtagandet inkluderas, i mån av tid och resurser hos Banverket, förutsatt att ingen annan trafik som redan omfattas av åtagandet

blir lidande. Detta illustreras med den streckade vägen i den övre delen av Figur 3 som tar en fungerande plan till en ny optimering via operatörernas önskemål.

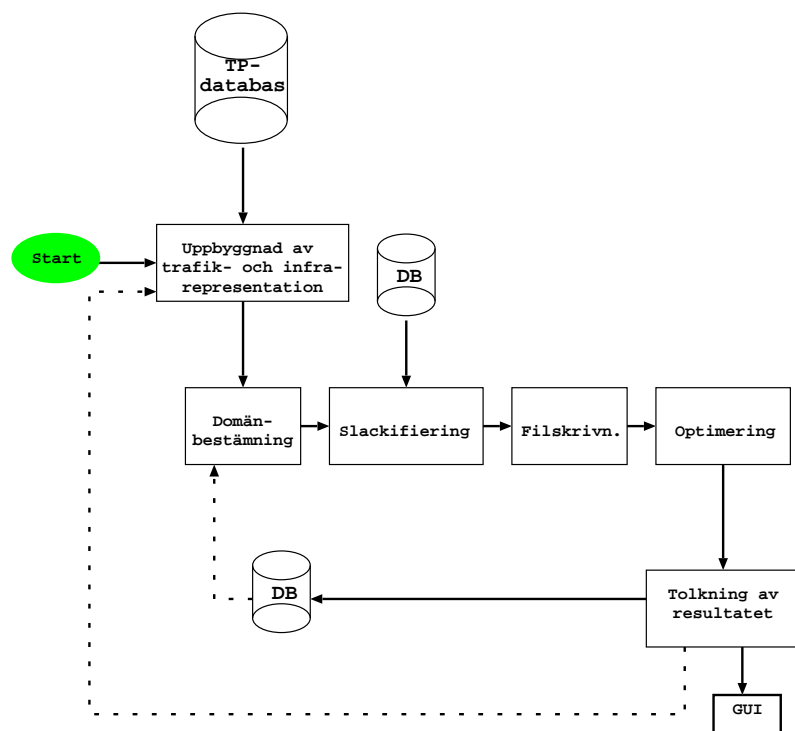
Det finns en viss risk att det inte finns någon produktionsplan som rymmer hela åtagandet. I detta fall får tidtabellägaren ta till de vanliga verktygen, dvs slackifiera för att identifiera möjliga åtgärder och sedan ta kontakt med berörda operatörer och diskutera olika lösningar. Att förlänga tågens gångtider, flytta på tåg samt stryka tåg borde i det här skedet krävas relativt sällan eftersom den grova planen i alla fall borde ha gett tillräckligt utrymme för trafiken i stora drag. Om sådana åtgärder blir vanliga och inte bara en sista utväg i enskilda och sällsynta fall, bör kriterierna för när den grova planen kan anses vara genomförbar ses över.

Oavsett hur en fungerande plan kommer till, dvs oavsett om det blev nödvändigt att modifiera åtagandet för att nå fram till en konfliktfri plan eller inte, finns det alltid en möjlighet att välja att fortsätta att optimera för att anpassa planen bättre till operatörernas övriga önskemål, trots att Banverket inte har åtagit sig att göra detta. Det är en iterativ process som när som helst kan avbrytas. Den färdiga produktionsplanen är då den bästa planen som uppfyller åtagandet som hittills hittats, eller den bästa planen som hittats som uppfyller det modifierade åtagandet i de fall det inte gick att hitta en produktionsplan som fullständigt uppfyllde det ursprungliga åtagandet. Målet är att den operativa tågledningen varje dygn ska ha bästa möjliga produktionsplan till sitt förfogande.

2.1.5 Det slutgiltiga åtagandet

När processen avslutats, meddelar Banverket Tågo och de andra operatörerna exakt vad de kan förvänta sig i form av tidpunkter för en delmängd av avgångarna och ankomsterna i produktionsplanen. Banverket kan lämna ut uppgifter om avgångar och ankomster som inte omfattas av åtagandet, men är i så fall noga med att understryka vilka delar av de kommunicerade produktionsplanerna som utgör åtagandet och vilka som inte gör det. Platser som definitivt utelämnas har övervägande produktionsteknisk karaktär, och hur de kommer att trafikeras kan komma att skilja från dag till dag.

Tidpunkterna som ingår i åtagandet är i de allra flesta fall löften om tidigaste avgång från respektive senaste ankomst till de olika platserna. Såvida operatören inte har haft specifika önskemål, tar sig Banverket friheten att under operativ drift leda fram tågen snabbare än planerat så länge tågen inte avgår före de utlovade tidpunkterna vid platser där åtagandet har specificerat en tidigaste avgång. Med andra ord ska ett tåg inte behöva lämna plattformen före annonserad avgångstid, men tågen kan i allmänhet räkna med att då och då komma fram före sin annonserade ankomsttid.



Figur 4: Demonstratorns olika delmoment

3 Demonstratorn

För att utreda vilket sorts datorstöd som lämpar sig till den nya tidtabellägningsprocessen, har vi utvecklat en demonstrator som inbegriper ett antal väldefinierade steg. Närmast följer en översiktlig beskrivning av hur processen i demonstratorn ser ut, där namnen på de olika stegen är explicit namngivna (se Figur 4). Resten av kapitlet ägnas åt att beskriva demonstratorn i detalj.

Figur 4 visar hur de olika delmomenten i processen hänger ihop. Utgångspunkten är alltid TP-databasen. Den innehåller grundförutsättningarna för trafiken, dvs information om tågen som ska köras, samt infrastrukturdatabas. Steget då en representation av trafiken och infrastrukturen byggs upp anpassar datamängden från TP-databasen till körningens ändamål. Under domänbestämningen konkretiseras ramarna för tågens avgång- och ankomsttider. Här väljer vi om data gällande tidpunkter och -intervall ska hämtas direkt från TP-databasen eller från någon annan källa. Därefter bestämmer vi i slackifieringssteget om vi vill tillåta lösningar där avgångs- och ankomsttider hamnar utanför de ursprungliga ramarna, och i så fall hur mycket tiderna får skilja. Detta steg kan om så önskas styras i relativt stor detalj. Enskilda tåg kan ges stora friheter medan andra låses helt. Vi kan också ange generella regler som att alla tåg får avgå upp till

ett bestämt antal sekunder tidigare eller senare än det ursprungligen var sagt.

I filskrivningssteget skrivs problemet ut på en fil på ett format som representerar optimeringsproblemets förutsättningar och som optimeringsverktyget CPLEX kan tolka. Därefter låter vi CPLEX läsa filen och genomföra en optimering i optimeringssteget. Resultatet sparas på en ny fil och kan om vi så väljer läsas in i en databas och användas i en senare körning, i så fall i domänbestämningssteget. Vi kan även välja att visualisera resultatet i ett grafiskt gränssnitt.

3.1 TP-databasen

Vår TP-databas baseras på TDEF-data, dvs data hämtat från TrainPlan. TDEF-data beskriver en tidtabell och infrastrukturen som tågen i tidtabellen ska köras på. Skapandet av TP-databasen från TDEF-data har utelämnats i arbetsgången som beskrivs i Figur 4 eftersom den bara behöver skapas en gång per studerat fall och därefter finns tillgänglig för alla framtida körningar på den aktuella trafiken. Hur TP-databasen skapas beskrivs inte i denna rapport.

Innehållet i TP-databasen sammanfattas i följande uppräkningslista och kompletteras med några förtydliganden i efterkommande stycken:

Om tågen

- vilka platser de passerar
- gångtider
- krav på uppehåll (plats och längd)
- avgångs- och ankomsttider

Om banan

- stationsegenskaper (t ex kapacitet)
- linjespårsegenskaper (t ex upp- eller nerspår)

Avgångs- och ankomsttiderna för tågen utgör tillsammans ett specifikt tidtabellutkast, dvs inte nödvändigtvis en fullständigt konfliktreglerad tidtabell. Med utkastets tider som utgångspunkt, kan vi testa olika optimeringsansatser för att försöka hitta lösningar som bättre motsvarar de ställda kraven. Det är emellertid inte nödvändigt att basera optimeringsproblemet direkt på dessa tider, vilket framgår nedan i beskrivningen av domänbestämningssteget.

Angående informationen om banan, dvs spårnätet, är dess konnektivitet endast implicit representerat. Hur platserna är förbundna framgår av hur tågen trafikerar nätet. En *gränsplats* är en plats som enligt den uppbyggda datastrukturen ser ut att vara förbunden med endast en annan plats i nätet. Även om det förekommer verkliga fall, såsom Malmö C och andra säckstationer, utgör de flesta gränsplatser tillsammans helt enkelt gränsen till de intilliggande tidtabellläggningsområdena.

3.2 Uppbyggnad av trafikrepresentationen

Grundidén i DDTP är att man ska planera på den nivå för vilken man har data. I detta steg hämtas en datamängd från TP-databasen som anpassas så att den matchar den aktuella detaljeringsnivån.

I ett tidigt planeringsstadium är det inte sannolikt att man behöver kunna uttala sig om tidpunkter för samtliga platser i nätet. Uppbyggnaden av en trafik- och infrastrukturepresentation genomförs enligt parametrar som avgör vilka av spårnätets platser som ska innefattas explicit av modellen, och vilka som ska döljas. Via dessa parametrar talar användaren om vilka platser som är *giltiga* för den aktuella detaljeringsnivån. Övriga platser *abstraheras bort*. Gränsplatser och platser där spårnätet förgrenar sig ska dock inte abstraheras bort. Inte heller platser där tåg startar och slutar (orsaken till det förklaras i avsnitt 3.3.1, “Indata”).

Den representation av spårnätet som byggs upp utgår från begreppet länk. En *länk* är en sträcka mellan två giltiga platser som inte innefattar någon annan giltig plats; en *abstrakt länk* är en länk som inbegriper minst en bortabstraherad plats. Information om vilka av länkarna som är abstrakta och hur många mötesstationer de innehåller är relevant, men i övrigt döljs de bortabstraherade platserna av modellen.

Samtidigt som länkrepresentationen av spårnätet byggs upp, skapas från TP-databasen en tågrepresentation som utgår från den nya spårnätsrepresentationen; tågen associeras med länkar i stället för med de enskilda spårsträckorna. Ett enskilt tågs väg i spårnätet modelleras av kedjor av två olika typer av aktiviteter: *platsaktiviteter* och *traverseringar*. Varannan aktivitet hos tåget är en traversering och varannan en platsaktivitet. Traverseringarna associeras med länkar och platsaktiviteterna med platser. Förekomsten av en platsaktivitet i kedjan av aktiviteter betyder att tåget passerar platsen i fråga, och eventuellt gör ett uppehåll där.

Nedan visas vilka egenskaper som förknippas med de två aktivitetstyperna. De egenskaper som är märkta med en asterisk (*) kan få sina värden utbytta senare i processen, medan övriga hör till tågens grundförutsättningar och betraktas som statiska i detta projekt.

Traversering

- avgångstid *
- traverseringstid
- headway
- tågdata (id, *business-id*, tågtyp)
- länkdata (id, riktning, antal bortabstraherade mötesstationer)

Platsaktivitet

- ankomsttid *
- uppehållslängd *
- typ
- tågdata (id, *business-id*)
- platsdata (namn, typ, kapacitet)

Traverseringstiden är tiden det tar för tåget att färdas längs länken. Om länken är abstrakt, dvs består av fler än en spårsträcka, är länkens traverseringstid summan av traverseringstiderna för de ingående spårsträckorna och de nödvändiga uppehållen på de bortabstraherade platserna.

Headway är ett mått på hur lång tid som måste passera mellan det att det aktuella tåget åker in på länken i fråga tills dess att nästa tåg i samma riktning får åka in på samma länk. Säkerhetsföreskrifterna är i själva verket långt mer komplicerade än så, men i detta sammanhang har vi bedömt denna approximation som tillräcklig. *Headway*-uppgifterna i vår TP-databas är baserade på blocksignalernas placering.

Kapaciteten hos en plats representeras med ett heltal och används för att beteckna hur många tåg platsen kan hysa samtidigt. Det är givetvis en förenkling av verkligheten eftersom den verkliga kapaciteten på många platser t ex beror på hur långa tågen är, växlarnas placering och konnektiviteten hos stationen.

Platsens typ säger vilken sorts plats det är fråga om, t ex blocksignal, krysstation, etc.

Business-id är det annonserade tågnumret.

3.3 Domänbestämning

Detta steg tillsammans med det efterföljande bestämmer vilka tidsramar de enskilda avgångarna och ankomsterna måste hålla sig till i den slutgiltiga lösningen, oavsett om målet är en grov plan eller en mycket detaljerad produktionsplan.

I det följande syftar *ursprungshandlingarna* på operatörernas ansökningar eller på Banverkets åtagande. Vilket som avses framgår av sammanhanget eftersom det beror på vilket syfte planeringen har. Om vi är i färd med att hitta en grov plan, syftar ursprungshandlingarna givetvis på operatörernas ansökningar. Om vi ska hitta en detaljerad produktionsplan i ett senare planeringskede, är "ursprungshandlingarna" helt enkelt åtagandet.

Uppdelningen i två steg är logisk. Tidsramarna som sätts i detta steg ska utgå från ursprungshandlingarna. Kraven på att tiderna ska hålla sig strikt innanför de tidsramar som beräknas i detta steg kan i sin tur vid behov relaxeras i steg två, slackifieringen.

Eftersom det är en iterativ process att ta fram åtagandet respektive detaljerade produktionsplaner, är det endast i den första iterationen som indata-materialet till detta steg är själva ursprungsansökningarna eller det ursprungliga åtagandet. I den mån det inte finns en godtagbar lösning som baseras på

ursprungshandlingarna, tar tidtabellkonstruktören hjälp av slackifieringssteget för att på ett kontrollerat sätt styra hur de ursprungliga kraven kan frångås och därmed identifiera hur operatörernas ansökningar eller Banverkets åtagande kan modifieras så att en så bra kompromiss som möjligt ska kunna hittas. Under följande iterationer, sker domänbestämningen på basis av de modifierade ursprungshandlingarna, tills en acceptabel lösning är funnen.

I brist på en databas över vad operatörerna ursprungligen ansökt om, utgår vi i stället från någon annan datamängd för att testa det program som vi har utvecklat. Vi kan t ex använda en delmängd av TP-databasen, i vilket fall datamängden motsvarar konturerna av ett verkligt tidtabellutkast från någon tidtabellkonstruktör, eller utgå från en av våra tidigare lösningar, under förutsättning att vi har sparat den på lämplig form (se sid 12).

3.3.1 Indata

Den datamängd vi väljer som representant för ursprungshandlingarna innehåller tidpunkter för ett urval av platserna i nätet. Det finns egentligen bara tre krav på indata till detta steg, och dessa ställs på tiderna som anges.

Tider måste

1. vara förenliga med tågens grundförutsättningar
2. gälla platser som är giltiga på aktuell abstraktionsnivå
3. anges åtminstone för platserna där tågen startar och slutar

Punkt ett betyder att tiderna som hämtas till detta steg inte får innebära att ett tåg ska traversera en länk snabbare än det är fysiskt möjligt. Vi har antagit att ett tågs gångtid inte kan understiga den som hämtas från TP-databasen¹, dvs en undre gräns för hur snabbt tåget kan gå sätts till summan av traverseringstiderna på de ingående spårsträckorna och de nödvändiga uppehållen på platserna som passeras (exklusive ett eventuellt uppehåll vid slutdestinationen).

Punkt två säger att de tider som anges måste gälla platser som är explicit representerade i vår modell. Rent praktiskt kan man givetvis använda ett indata-material som även anger tider för bortabstraherade platser, men dessa kommer att kastas bort eftersom platserna i fråga inte "finns" (syns) i programmet.

Punkt tre är ett krav av mer filosofisk (men även implementationsteknisk) karaktär: för att kunna säga någonting om tåget överhuvudtaget, behöver vi veta åtminstone när det ska börja och när det ska sluta. Tillsammans med

¹Det kan tyckas självklart vid en första anblick, men egentligen finns det faktiskt inget som säger att traverseringstiden i TP-databasen är en undre gräns för gångtiden på sträckan för tåget i fråga. Traverseringstiden i TP-databasen är helt enkelt den som tidtabellkonstruktören hade tänkt sig för detta tåg i det specifika tidtabellutkast som TP-databasen är skapat utifrån. Den kan inbegripa alla möjliga tillägg. Även om det ur TDEF-data går att urskilja vad som är grundtraverseringstiden på en sträcka och vad som är tillägg, har vi i den här delen av projektet ansett att traverseringstiderna i tidtabellutkastet är tillräckligt goda approximationer av gångtiderna för vårt syfte.

grundförutsättningarna kan vi i så fall bestämma ramarna för hur tåget kommer att bete sig på alla mellanliggande platser.

Vad gäller tiderna, kan de vara angivna antingen som exakta klockslag eller som intervall, s k *domäner*. Med detta vill vi efterlikna att en operatör t ex har uttryckt önskemål om att en viss avgång ska ske någon gång mellan 13:24 och 13:27, eller att ett uppehåll ska vara mellan 2 och 3 minuter långt. Om en ankomst är angiven utan domän, förutsätts att operatören godkänner att tåget eventuellt ankommer tidigare, och för avgångar gäller tolkningen att en senare avgång är acceptabel – så länge som den efterföljande ankomsten sker i tid.

3.3.2 Beräkning av tider

I detta steg konkretiseras ramarna för en delvis specificerad plan, i syfte att hitta en mer detaljerad plan som överensstämmer med den givna specifikationen. Med andra ord görs planeringsutrymmet explicit för ett givet antal förutsättningar (t ex operatörernas ansökningar) så att vi i första hand ska kunna hitta en mer detaljerad plan inom denna ram. Notera att det inte är säkert att det finns någon konkret lösning som följer specifikationen, och att vi i så fall måste relaxera kraven (se avnitt 3.4).

Vi antar här att vi utgår från de exakta ankomst- och avgångstider som i ursprungshandlingarna anges för ett urval av platserna som ingår i tågets väg². Från dessa vill vi beräkna därmed förenliga tidsintervall för de ankomster och avgångar som inte är helt bestämda, på alla platser som är explicit representerade på den aktuella detaljeringsnivån. Indata till beräkningen är information om de aktiviteter som ska utföras enligt den konkretiserade planen. Vi förutsätter att vi också har tillgång till gångtiderna för alla mellanliggande delsträckor och att dessa är fixa.

Låt T_a beteckna tiden det tar mellan ankomsterna vid två på varandra följande platser där ursprungshandlingarna (ansökningarna eller åtagandet) specificerar en önskad tidpunkt, hädanefter kallade *relevanta* platser. Tiden för sträckan mellan två relevanta platser innefattar i allmänhet tidspåslag för t ex förväntade möten och förbigångar. Summan av gångtiderna på de mellanliggande sträckorna och de kortaste möjliga uppehållen på sträckans platser utgör däremot den kortaste möjliga tiden för hela den sammansatta sträckan utan hänsyn till annan trafik. Skillnaden mellan dessa båda tidsutsträckningar utgör det slack vi har att fördela över uppehållen på sträckans platser. Observera att vi med sträckans platser även avser platsen som tåget avgår från när det ska traversera sträckan, medan platsen som traverseringen av sträckan avslutas på inte inkluderas.

För varje par av två på varandra följande relevanta platser, låt A_0 och A_n representera deras helt bestämda ankomsttider, medan A_1, \dots, A_{n-1} får mot-

²Det är egentligen inget som hindrar att vi istället utgår från intervall även för avgångs- och ankomsttider på de för tågen utvalda platserna och/eller för gångtiderna på de sträckor som ligger mellan dessa platser. Den mekanism som beskrivs här kan relativt enkelt generaliseras för att hantera även dessa fall men då det komplicerar framställningen väljer vi att endast beskriva fallet för fixa tider på de platser som valts ut för tågen och området i fråga.

svara ankomsttiderna på de $n - 1$ mellanliggande platserna, vilka ska beräknas. Låt på samma sätt D_0, \dots, D_n representera motsvarande avgångstider. Uppehållstiderna S_0, \dots, S_{n-1} på platserna ges då av $S_i = D_i - A_i$ för alla $0 \leq i < n$. Låt oss anta att den undre gränsen \underline{S}_i för uppehållstiden för alla $0 \leq i < n$ är känd, och att likaså traverseringstiden $T_i = A_i - D_{i-1}$ för alla $0 < i \leq n$ är känd.

I detta sammanhang behöver vi påpeka att en "helt bestämd" ankomsttid ska tolkas som "ankommer ej senare än", medan ett krav på avgångstid innebär "avgår ej tidigare än". Detta innebär att vi anser att ett tåg inte bryter mot något krav om det ankommer tidigare än den önskade ankomsttiden eller avgår senare än den önskade avgångstiden – så länge det inte missar senast ankomstnågonstans under sin färd, eller avgår tidigare än det får från någon plats.

Låt nu

$$\underline{T} = \left(\sum_{i=1}^n T_i + \underline{S}_i \right) - S_n$$

vara den kortaste möjliga tiden för att traversera hela sträckan mellan de två relevanta platserna. Låt vidare $V = T_a - \underline{T} - \underline{S}_0$ vara det totala tillgängliga slacket på samma sträcka. Då vi inte vet något om hur det faktiskt kommer att fördelas är den mest allmänna slutsatsen att $\underline{S}_i \leq S_i \leq V + \underline{S}_i$ för alla $0 \leq i < n$.

Vi låter $\underline{D}_0 = A_0 + \underline{S}_0$ och $\overline{D}_0 = \underline{D}_0 + V$, där A_0 är det värde som hämtas från ursprungshandlingarna och representerar önskad ankomst till den relevanta platsen i fråga. I så fall fås villkor för hur avgångs- och ankomsttider kan variera av följande uttryck som i vår kod implementerats som en enkel iteration över de mellanliggande platserna.

$$\begin{aligned} \underline{D}_k &= \underline{D}_0 + \sum_{i=1}^k T_i + \underline{S}_i \\ \underline{D}_k &= \underline{D}_k + V \\ \underline{A}_k &= \underline{D}_{k-1} + T_k \\ \underline{A}_k &= \underline{D}_{k-1} + T_k \end{aligned} \quad \text{för alla } 0 < k < n$$

där \underline{X} representerar den undre, och \overline{X} den övre gränsen för en variabel X .

Denna iteration utförs för varje par av på varandra följande relevanta platser i tågets väg, där en relevant plats alltså definieras som en plats för vilken det finns en specifik tid som ska upprätthållas angiven i indata materialet.

3.4 Slackifiering

Direkt efter domänbestämningssteget har vi fullständiga data för att genomföra en optimering, och det är också rimligt att inleda med att pröva detta. Först om resultatet inte blir tillfredsställande vill vi börja experimentera med att ge beräkningsverktyget viss frihet att frångå de ramar som domänbestämningssteget har gett oss. Konkret sker detta genom att datamängden *slackifieras*.

Termen slackifiering kommer från engelskans *slack*, i betydelsen "en del som är tillgänglig men inte utnyttjad"³. Att slackifiera en datamängd från domänbe-

³"A part that is available but not used" (<http://www.merriam-webster.com/dictionary/slack>).

stämningssteget innebär att öka storleken på några eller alla domäner med syfte att möjliggöra en eller flera lösningar. Att flytta domäner ingår däremot inte i slackifieringen utan normalt ska den gamla domänen helt rymmas inom den nya. För att återknyta till exemplet i föregående avsnitt, kan operatörens önskemål om en avgång mellan 13:24 och 13:27 i detta steg ändras till att avgången i stället får ske t ex mellan 13:21 och 13:29.

Även om det finns möjlighet att detaljstyra acceptabla tidpunkter för varje enskild avgångs- och ankomsttid, är det givetvis tidsödande att göra så och oftast inte heller meningsfullt eftersom resultatet av slackifieringen i första hand ska vara *vägledande* och identifiera vilka åtgärder som *skulle kunna vidtas* för att en bra kompromiss ska kunna nås.

Slackifiering sker i huvudsak enligt de övergripande principer som strax ska beskrivas. Detaljerade krav kan man därefter lägga till för att inskränka ramarna i de fall den generella slackifieringen varit för generös och gjort det möjligt för en lösning att bryta mot krav som man fortfarande bedömer som nödvändiga att upprätthålla.

3.4.1 Generella principer

Vi skiljer på två olika parametrar när vi slackifierar. Den ena styr hur vi kan tänka oss att förlänga uppehållen. Den andra parametern styr hur mycket tågets avgång respektive ankomst får skilja sig jämfört med de givna ramarna från domänbestämningsteget. Notera att denna andra parameter endast detaljreglerar tidpunkterna där tåget startar och slutar, närmare bestämt tidigaste avgång vid tågets första plats och senaste ankomst till tågets sista plats⁴.

Om det inte finns någon lösning inom de ramar som domänbestämningsteget har fastställt, får vi ställa in oss på att hitta en kompromisslösning. För att lyckas hitta en sådan lösning måste vi i många fall sannolikt släppa på kraven för tidigaste avgång respektive senaste ankomst för vissa relevanta platser⁵ och i värsta fall även för tågets första och sista plats. Motiveringen till att ha en parameter som styr vad som sker just vid tågens första och sista platser är att dessa ofta sammanfaller med gränsplatser som per definition gränsar till andra tidtabellkonstruktionsområden⁶. Att frånga ramarna på gränsplatser är extra känsligt eftersom det sannolikt innebär att lösningarna för de olika tidtabellkonstruktionsområdena inte längre fungerar ihop.

Hur slackifieringen ska förhålla sig till platser som inte är tågens första och sista platser bestäms individuellt per tåg och plats av extra indata. Om inget särskilt anges, ignoreras alla krav på tidigaste avgång och senaste ankomst vid sådana platser och slackifieringen fokuserar på att tåget som helhet ska avgå och ankomma inom de tillåtna intervallen vid tågens startplatser och slutdestinationer.

⁴ Senaste möjliga avgång från första platsen och tidigaste möjliga ankomst till sista platsen för tåget påverkas förvisso även av denna parameter, men styrs inte lika hårt då även andra aspekter spelar in.

⁵ I 3.3.2 definierar vi begreppet relevant plats.

⁶ Undantaget är gränsplatser som i själva verket är säckstationer. Jämför med resonemanget om gränsplatser i 3.1.

3.4.2 Beräkning

Vi delar upp slackifieringen i två steg. I det första utvidgas avgångs-, ankomst- och uppehållstidsintervall utifrån givna globala parametrar, och i det andra inskränks eventuellt vissa av dessa vidgade intervall baserat på lokala (plats-specifika) villkor.

3.4.3 Allmän utvidgning

Den allmänna utvidgningen av domänerna i ett indatamaterial (t ex utdata från domänbestämningssteget ovan) styrs av två parametrar L_s och L_e som anger hur varje uppehåll får förlängas i förhållande till sin minimitid, respektive specificerar hur intervallen till den första avgångstiden och den sista ankomsttiden i tågets resa ska utvidgas.

Om vi betraktar endast tågets första och sista platser som *relevanta*, kan vi på samma sätt som i avsnitt 3.3.2 låta D_0, \dots, D_n och A_0, \dots, A_n representera avgångs- respektive ankomsttiderna på tågets resa. Antag vidare att T_1, \dots, T_n och S_0, \dots, S_n representerar traverserings- respektive uppehållstider på alla sträckor och (för abstraktionsnivån giltiga) platser i tågets väg. Låt oss också anta att $S_i = D_i - A_i$ gäller för alla $0 \leq i \leq n$ och $T_i = A_i - D_{i-1}$ gäller för alla $0 < i \leq n$. Låt vidare \underline{X} representera den undre, och \overline{X} den övre gränsen för varje variabel X . Från domänbestämningssteget har vi också initiala domäner $[\underline{X}, \overline{X}]$ som uppfyller villkoren ovan för alla variabler.

Vi vill nu bestämma en utvidgning av dessa intervall givet de två parametrarna L_s och L_e .

3.4.4 Algoritm

Algoritmen för att beräkna utvidgningen sveper över sekvensen av avgångs-, uppehålls- och ankomsttider och beräknar nya gränser för D_i och A_i samt en eventuell utvidgning av den övre gränsen för S_i för alla $0 \leq i \leq n$, utgående från en första avgångstid D_0 , en sista ankomsttid A_n , de två angivna parametrarna L_s och L_e , samt de initiala gränserna för S_i .

I algoritmen nedan betecknas variabeln D_i med $D[i]$, A_i med $A[i]$, S_i med $S[i]$ medan gränsen $\underline{D_i}$ betecknas med $Dmin[i]$, och $\overline{D_i}$ med $Dmax[i]$ samt motsvarande för A_i och S_i . Parametrarna L_e och L_s betecknas med Le respektive Ls . Kommentarer i koden föregås av #.

```
Dmin[0] = D[0] - Le; # Undre gräns för första avgången
for (i = 1; i++; i < n) do
  Amin[i] = Dmin[i-1] + T[i]; # Undre gräns för A[i]
  Dmin[i] = Amin[i] + Smin[i]; # Undre gräns för D[i]
end
Amin[n] = Dmin[n-1] + T[n]; # Undre gräns för sista ankomsten
Amax[n] = A[n] + Le; # Övre gräns för sista ankomsten
for (i= n-1; i--; i > 0) do
  Dmax[i] = Amax[i+1] - T[i+1]; # Övre gräns för D[i]
  Amax[i] = Dmax[i] - Smin[i]; # Övre gräns för A[i]
```

```

    Smax[i] = max(Smax[i], Smin[i] + Ls); # Relaxera Smax[i]
end
Dmax[0] = Amax[1] - T[1]; # Övre gräns för första avgången
Smax[0] = max(Smax[0], Smin[0] + Ls); # Relaxera Smax[0]

```

Notera att L_s och $\overline{S_i}$ i denna beräkning inte påverkar gränserna för D_i och A_i . Detta innebär dock inte att dessa parametrar är oberoende, bara att detta beroende inte utnyttjas i denna fas av planeringen.

3.4.5 Specifik inskränkning

De specifika inskränkningarna i slackifieringen ger oss möjlighet att inskränka hela tågets schema baserat på lokala begränsningar på avgångs-, ankomst- och uppehållstider. Dessa lokala begränsningar tar formen av intervall inom vilka den berörda tidpunkten eller tidsutsträckningen ska falla. Dessa inskränkningar påverkar dock inte bara vad som händer lokalt på den plats variabeln associeras med, utan via villkoren $S_i = D_i - A_i$ och $T_i = A_i - D_{i-1}$ även tidigare och senare händelser på tågets väg.

Mekanismen för att beräkna de icke-lokala konsekvenserna av sådana inskränkningar är baserad på s.k. villkorspropagering men påminner mycket om algoritmen som beskrivs i avsnitt 3.4.4 ovan med den skillnaden att man kan behöva svepa över sekvensen av tider och aktiviteter både framåt och bakåt i tiden och ta hänsyn till både undre och övre gränser för uppehållstiderna innan alla konsekvenser beräknats.

Så kan t ex en inskränkning av den övre gränsen för en ankomsttid potentiellt påverka inte bara den övre gränsen för uppehållet på den aktuella platsen, utan också de övre gränserna för både tidigare och senare händelser på tågets väg. Liknande konsekvenser kan uppstå från inskränkningar av både undre och övre gränser för uppehållstider.

Vi ska här inte gå in på alla detaljer i hur denna beräkning sker utan nöja oss med att konstatera att vi i villkorspropagering för linjära likheter som $S_i = D_i - A_i$ och $T_i = A_i - D_{i-1}$ har en effektiv metod att härleda alla konsekvenser av inskränkningar på de berörda variablerna. Vi kan också notera att detta gäller även för mer komplicerade inskränkningar av variablernas domäner än intervall men att mekanismerna för att göra detta är mest effektiva så länge vi håller oss till intervall.

I den implementering vi har finns för närvarande ingen fullständig mekanism för att garantera att alla de angivna specifika inskränkningarna är förenliga med den allmänna utvidgningen eller med varandra. Villkor som inte är det genererar visserligen felmeddelanden och ignoreras men denna mekanism är, i fallet att villkoren är sinsemellan motsägande, beroende av i vilken ordning villkoren behandlas. Ett sådant felmeddelande ska alltså i allmänhet tolkas som att hela mängden av lokala villkor är motsägelsefulla och inte att det specifika villkor som anges i felmeddelandet nödvändigtvis i sig bär hela skulden till att inskränkningen som helhet misslyckats.

3.5 Filskrivning

I filskrivningsteget skrivs problemet ut på en fil på ett format som representerar optimeringsproblemets förutsättningar och som optimeringsverktyget CPLEX kan tolka. Det här avsnittet beskriver de grundläggande principerna för problemrepresentationen. För en detaljerad redogörelse hänvisar vi till [1], "Slutrapport för projektet TUFF, TågplaneUtveckling För Framtiden".

3.5.1 Terminologi och beteckningar

Ett tåg modelleras av kedjor av alternerande platsaktiviteter och traverseringar, där traverseringarna associeras med länkar och platsaktiviteterna med platser. Mer specifikt associeras avgångar med länkar medan ankomster och uppehåll associeras med platser. Följaktligen avgår ett tåg *ut på en länk* medan det ankommer *till en plats* och har sitt eventuella uppehåll *på* platsen.

Följande beteckningar används:

Trn index för ett tåg

Lnk index för en länk

Lcn index för en plats

dep tidpunkten för en avgång (ut på en länk)

arr tidpunkten för en ankomst (till en plats)

tt durationen för en traversering (på en länk)

w durationen för ett uppehåll (på en plats)

hw headway

3.5.2 Tågekvationer

Grundkraven på aktiviteterna hos ett enskilt tåg är att de ska vara i tidsmässig sekvens vilket modelleras med hjälp av ekvationer som vi kallar *tågekvationer*: ett tåg måste ankomma till en plats innan det kan avgå därifrån. Om tåget har ett uppehåll på platsen i fråga, är längden på uppehållet skillnaden mellan tågets avgångstid från och ankomsttid till platsen. Om en platsaktivitet hos tåget inte har något uppehåll förknippat med sig, ankommer och avgår tåget på samma klockslag till/från en viss plats och uppehållstiden, $w_{Trn,Lcn}$, är 0.

Tågekvationerna motsvarar i princip kedjorna av alternerande platsaktiviteter och traverseringar där ett tågs färd beskrivs med ankomster till platser och avgångar ut på länkar. Varje plats som passeras av tåget, utom tågets första och sista plats, är förknippad med precis två ekvationer, vars principiella konstruktion illustreras nedan:

$$\begin{cases} dep_{Trn,Lnk2} = dep_{Trn,Lnk1} + tt_{Trn,Lnk1} + w_{Trn,Lcn} \\ tt_{Trn,Lnk1} = arr_{Trn,Lcn} - dep_{Trn,Lnk1} \end{cases}$$

Vi illustrerar några aspekter av tågekvationerna med hjälp av ett utdrag ur en av våra LP-filer från en körning som visar samtliga tågekvationer för ett specifikt tåg:

$$\begin{aligned}
d_{9_24} + w_{9_16} - d_{9_19} + tt_{9_24} &= 0 \\
a_{9_16} - d_{9_24} - tt_{9_24} &= 0 \\
d_{9_19} + w_{9_13} - d_{9_21} + tt_{9_19} &= 0 \\
a_{9_13} - d_{9_19} - tt_{9_19} &= 0 \\
d_{9_21} + w_{9_22} - d_{9_29} + tt_{9_21} &= 0 \\
a_{9_22} - d_{9_21} - tt_{9_21} &= 0 \\
a_{9_24} - d_{9_29} - tt_{9_29} &= 0
\end{aligned}$$

Alla tåg, platser och länkar får index under en körning för att variablerna ska få unika namn som CPLEX kan tolka. Eftersom avgångar och traverseringar är förknippade enbart med platser och ankomster och uppehållsvariabler med länkar, kan en plats och en länk få samma index utan risk för sammanblandning.

Konkret passeras följande platser av tåget, som har fått index 9 i körningen (platsindex inom parentes): LÅ(14), LÅÖ(16), LLN(13), VT(22) och ÖJ(24). Mellan platserna traverserar tåget följande länkar (länkindex inom parentes): LÅ-LÅÖ(24), LÅÖ-LLN(19), LLN-VT(21) och VT-ÖJ(29).

Det finns ingen orsak att titta på utdraget i mer detalj än att vi kan göra några konkreta observationer angående tågets första och sista platser. Tåget börjar i LÅ och slutar i ÖJ, platser som har fått index 14 respektive 24, men det finns inte vare sig någon ankomstvariabel eller någon uppehållsvariabel för tåget i LÅ. Inte heller finns det någon uppehållsvariabel i ÖJ eller någon avgångsvariabel för avgången från ÖJ. Där tåget startar, har vi tagit det modelleringsmässiga beslutet att bara bry oss om avgången, och där tåget slutar är bara ankomsten intressant. Detta överensstämmer med praxis i TrainPlan, där start- och slutplatsernas uppehållstid alltid är noll.

3.5.3 Länkekvationer

Varje länk i vår modell är en begränsad resurs och hur den kan utnyttjas specificeras med hjälp av länkekvationer. Principen lyder att två tåg inte får befinna sig på samma länk samtidigt om de är motriktade, och att de i det medriktade fallet i varje ögonblick måste upprätthålla ett tidsmässigt avstånd från varandra som minst motsvarar headway.

Det vi kallar *länkekvationer* modellerar kravet att en avgång ut på en länk sker på ett säkert avstånd från alla andra avgångar ut på samma länk. I det motriktade fallet måste vi kräva att avgången för det ena tåget sker minst så långt innan det andra tågets avgång att det första tåget hinner traversera länken innan det andra tåget ger sig ut på samma länk.

Med andra ord måste precis en av ekvationerna nedan vara uppfylld för varje länk och par av tåg som är motriktade och potentiellt kan mötas på länken:

$$\begin{cases} dep_{Trn1,Lnk} > dep_{Trn2,Lnk} + tt_{Trn2,Lnk} \\ dep_{Trn2,Lnk} > dep_{Trn1,Lnk} + tt_{Trn1,Lnk} \end{cases}$$

När det gäller medriktade traverseringar på länkar, måste headway mellan tågen upprätthållas vid såväl länkens avgångs- som ankomstplats. Dessutom måste vi kräva att tågen ankommer vid länkens ankomstplats i samma ordning som de avgick från länkens avgångsplats. Om vi inte ställer det sistnämnda kravet, tillåter vi i praktiken att det ena tåget kör om det andra (förutsatt att ankomster och avgångar separeras tillräckligt i tiden), vilket vi naturligtvis inte kan godkänna på ett enkelspår utan mötesstationer.

Således, för det medriktade fallet ska vi kräva att precis ett av följande par av ekvationer är uppfyllt, där L_{cn} är platsen båda tågen ankommer till när de har traverserat den gemensamma länken L_{nk} :

$$\begin{cases} dep_{Trn1,Lnk} > dep_{Trn2,Lnk} + hw_{Trn2,Lnk} \\ arr_{Trn1,Lcn} > arr_{Trn2,Lcn} + hw_{Trn2,Lnk} \end{cases}$$

eller

$$\begin{cases} dep_{Trn2,Lnk} > dep_{Trn1,Lnk} + hw_{Trn1,Lnk} \\ arr_{Trn2,Lcn} > arr_{Trn1,Lcn} + hw_{Trn1,Lnk} \end{cases}$$

Notera att vi endast skriver ut ekvationer eller ekvationsgrupper för de par av tåg vars domäner ligger tillräckligt nära varandra i tiden för att de ska kunna behöva länken i fråga samtidigt.

Under själva problemformuleringen tas inget beslut om vilka av de parvis uteslutande ekvationer eller ekvationsgrupper som ska vara uppfyllda, eftersom det skulle innebära samma sak som att vi på förhand bestämmer vilket av två tåg som har företräde på en viss länk. Detta är naturligtvis en av de saker som CPLEX ska avgöra åt oss. Samtidigt kan vi inte ge båda ekvationer eller ekvationsgrupper såsom de är formulerade ovan till CPLEX eftersom de omöjligt kan vara uppfyllda samtidigt: CPLEX kommer bara att svara att det inte finns någon lösning.

Vi använder därför något som inom heltalsprogrammering kallas stora M-metoden (Big M Method). Den går ut på att man inför binära variabler på ett sådant sätt att beslutet om vilka ekvationer som ska vara uppfyllda tas av lösaren, dvs CPLEX i vårt fall, medan den söker efter en lösning. Detaljer om hur stora M-metoden fungerar och hur vi använder den återfinns i [1].

3.5.4 Platsekvationer

I vår modell är en plats ett punktformat objekt. I stället för att hålla reda på de olika spår som platsen inbegriper i verkligheten, har varje plats en *kapacitet* som motsvarar det antal tåg platsen förväntas kunna hysa samtidigt. Detta är givetvis en grov förenkling, men ger oss samtidigt ett relativt billigt sätt (beräkningstidsmässigt) att kontrollera att stationerna inte belastas på ett alldeles orimligt sätt.

Punktformigheten innebär också att en platsaktivitet börjar i samma ögonblick som tåget ankommer till platsen och avslutas samtidigt som tåget avgår från platsen med avseende på kraven på att aktiviteterna ska ske i sekvens. Med

avseende på stationsbeläggning anses dock även en platsaktivitet utan uppehåll ha en viss symbolisk utsträckning eftersom platsens kapacitet måste tas i beaktan.

Platsekvationerna uttrycker att beläggningen av stationerna måste ske på ett sådant sätt att platsens kapacitet inte överstigs. Tåg som potentiellt överlappar varandra på platsen måste med andra ord sekvensieras så att kapaciteten respekteras. Hur dessa villkor kan formuleras på ett effektivt sätt är ganska komplicerat att beskriva och finns redan förklarat i [2].

3.5.5 Styv tidtabell och liknande krav

Med *styv tidtabell* menas sådana typer av krav som säger att tåg ska avgå eller ankomma med vissa bestämda tidsmässiga avstånd från varandra, vid samtliga tidtabellagda platser som de trafikerar. Även om begreppet styv tidtabell normalt endast används för trafikupplägg där de inblandade tågen ligger nära varandra i tiden, t ex inom en timme, är det även vanligt förekommande med liknande krav på trafiken där det tidsmässiga avståndet mellan tågen är mycket större än så. T ex förväntas tåg med samma tågnummer gå likadant varje dag.

Krav på att vissa avgångar och ankomster ska tidtabelläggas på ett bestämt avstånd från varandra – oavsett om avståndet är en kvart eller 24 timmar – är enkelt att uttrycka i LP-filen, men kompliceras något av att en styv tidtabell (och andra liknande krav) inte nödvändigtvis innebär att tågen som ingår i det styva upplägget behöver följa varandra identiskt i praktiken. En *annonserad avgång* är nämligen inte samma sak som en *verklig avgång*, och så länge den verkliga avgången inte planeras ske tidigare än den annonserade avgången, kan de representera två skilda tidpunkter utan att det innebär något problem. Den annonserade tiden är helt enkelt den som kunden ser i tidtabellen och den andra är den som tågledaren ser när han eller hon ska leda fram tåget. Motsvarande gäller naturligtvis för ankomster, där en *annonserad ankomst* kan gälla en senare tidpunkt än den tid som planeras för den verkliga ankomsten.

Det är de verkliga ankomsterna och avgångarna som ingår i tåg-, länk- och platsekvationerna. För att upprätthålla eventuella krav på styv tidtabell, införs variabler även för annonserade ankomster och avgångar. Närhelst två avgångar ska ske med ett bestämt tidsmässigt mellanrum, ställs det kravet endast på de annonserade avgångarna, inte på de verkliga avgångarna. Motsvarande gäller förstas ankomster. Samtidigt ser vi till att en annonserad avgång sker senare än (eller samtidigt som) den verkliga avgången, och att en annonserad ankomst sker senast när den verkliga sker, genom att introducera ekvationer för dessa krav.

3.5.6 Domäner

Förutom ekvationer som ska uppfyllas, innehåller LP-filen domäner (eng. bounds), för alla eller vissa av variablerna. Om en domän är specificerad för en variabel, måste värdet för variabeln ligga inom domänen för att lösningen ska vara giltig (eng. feasible). Om en variabel inte är förknippad med någon domän, kan

CPLEX däremot tilldela vilket värde som helst till variabeln så länge det uppfyller alla andra krav som ställs på den.

3.5.7 Binärdeklaration

Ekvationerna så som de har beskrivits i avsnitten 3.5.2-3.5.4 utgör tillsammans med domäner och *binärdeklarationer* ett s.k. MIP-problem. MIP står för Mixed Integer Programming och kan översättas till blandad heltalsprogrammering. I ett MIP-problem är det till skillnad från i ren heltalsprogrammering endast vissa variabler som måste vara heltal.

I vårt fall är det de binära variablerna som vi bl.a. introducerar i samband med stora M-metoden som vi kräver ska vara heltal, och då specifikt antingen heltalet 0 eller 1. För att de ska uppfylla detta krav, måste vi binärdeklarera dem explicit i LP-filen.

3.5.8 Objektfunktion

Utan objektfunktion, kommer alla tilldelningar av värden på variabler som ger en giltig lösning att vara likvärdiga. Om vi däremot inför en objektfunktion som är kopplad till någon eller några av våra variabler, kommer lösaren (CPLEX) att optimera för att *minimera kostnaden* för lösningen.

En sak är särskilt viktig att poängtera när det gäller kostnaden, och det är att den inte ska ses som mätt i kronor. Objektfunktionen kan vara vilken ekvation som helst som ska minimeras. T.ex. skulle vi kunna ange en optimeringsfunktion som summan av alla tågs uppehållstider och därmed få en lösning som innehåller den minsta möjliga sammanlagda uppehållstiden – förutsatt att det överhuvudtaget finns en lösning på vårt problem.

Om det finns flera aspekter att ta hänsyn till vad gäller optimering, är det en avvägning hur man vill vikta kostnaderna för de olika aspekterna mot varandra. Upphållstid kanske inte ska betraktas som lika illa för lägre prioriterade tåg som för högre prioriterade tåg, bara för att nämna ett exempel. För att styra lösningarna mot att ge lågt prioriterade tåg uppehållstid hellre än de högt prioriterade tågen, kan man i objektfunktionen multiplicera persontågens uppehåll med någon faktor för att göra sådana uppehåll dyrare.

En beskrivning av hur vi har använt objektfunktionen återfinns i närmast följande avsnitt.

3.6 Optimering

Slutmålet för processen är en optimal plan. En optimal plan betyder i det här sammanhanget dels optimal i den intuitiva bemärkelsen “så bra som möjlig”, dels optimal från ett strikt matematiskt perspektiv. De två hänger ihop eftersom vi bara kan få en optimeringsprogramvara att ge oss en så bra plan som möjligt om vi definierar optimalitet på ett intuitivt och samtidigt korrekt matematiskt sätt.

Det är viktigt att skilja på vad som är en lösning på vårt problem och vad som är en lösning från ett optimeringsverktyg. Något förenklat uttryckt är vårt specifika problem löst när vi har tagit fram en fungerande produktionsplan, vilket i praktiken är detsamma som en sekundupplöst, konfliktreglerad tidtabell. För att komma fram till en lösning på vårt problem, använder vi oss däremot av en eller flera körningar i ett optimeringsverktyg. Vi använder uteslutande optimeringsverktyget CPLEX för våra optimeringar och hänvisar i fortsättningen till denna programvara, men följande är givetvis sant även för andra, liknande programvaror.

En körning i CPLEX resulterar antingen i att programmet hittar en lösning, terminerar med meddelandet att det inte finns någon lösning, eller att programmet misslyckas både med att hitta en lösning och att meddela att det inte finns någon lösning inom en rimlig tidsrymd. Konkret har vi en sk giltig lösning (eng. feasible solution) omedelbart när vi har hittat en uppsättning tidpunkter för alla avgångar och ankomster som tillsammans uppfyller de krav som finns uttryckta i LP-filen. Om lösningen är optimal eller inte beror dels på hur vi har definierat optimalitet, dels på om CPLEX har fått arbeta klart eller om körningen blev avbruten.

Beroende på vilka krav vi ställer på trafiken är det inte säkert att det finns en lösning. I så fall är ett alternativ att slackifiera – eventuellt i olika omgångar med t ex successivt större slack tills det finns en lösning för CPLEX att hitta. En lösning till det slackifierade problemet är förvisso ingen lösning till det ursprungliga problemet. Men en sådan fanns ju inte, så lösningen är en plan som förhoppningsvis ligger nära det man ville ha.

Stora slack ger i regel lång exekveringstid eftersom antalet kombinationer av tidpunkter som ska utvärderas ökar ju större slack man inför. Dessutom riskerar tåglägen att hamna långt från sina utgångslägen⁷, vilket sällan är önskvärt. Om möjligt, bör problemet i stället modelleras på ett sådant sätt att det går att identifiera vad i kraven som är omöjligt att uppfylla. Vägled av resultatet av en sådan optimering, kan tidtabelläggaren avgöra vilka riktade insatser som verkar nödvändiga. Dessa kan röra slack på enskilda tåg eller grupper av tåg, respektive flytt eller strykning av tåg.

Modellering av problemet på ovannämnda sätt går ut på att man släpper på vissa typer av krav så att brott mot dessa är tillåtna men förknippade med straff i form av en kostnad. Kostnaden är, som nämnt tidigare, abstrakt och har ingenting med pengar att göra. Ett optimeringsverktyg såsom CPLEX kan därefter optimera mot att minimera en sådan kostnad. Effekten blir att en lösning innehåller den billigaste delmängden av brott mot kraven som möjligt⁸. Den är med andra ord optimal utifrån de krav man har ställt.

Det viktigaste och mest konkreta exemplet från vår ansats är att vi kan välja att tillfälligt släppa på det absoluta kravet att en enkelspårsträcka bara får hysa

⁷ Detta är sant under förutsättning att man inte optimerar mot att komma så nära de ursprungliga kraven som möjligt. Det sistnämnda bedömer vi efter inledande test som dyrt räknat i komplexitet och exekveringstid.

⁸ Om CPLEX blir avbruten i förtid returneras lösningen med den lägsta kostnaden bland de lösningar som har hittats hittills.

ett tåg i taget. En lösning kommer i så fall kunna innehålla sådana resurskonflikter, men det kan alltså vara värdefullt i sig att se var sådana konflikter uppstår, och det är därför vi tillfälligt vill tillåta dem. Notera att om vi tillåter konflikter mot en kostnad i en körning, kommer en konfliktfri lösning på problemet att ha kostnad noll: förutsatt att en lösning finns och CPLEX kan hitta den optimala lösningen inom den tid som vi ger den att jobba med optimeringen, innehåller lösningen som CPLEX returnerar bara konflikter om det överhuvudtaget inte finns en konfliktfri lösning.

CPLEX ges möjligheten att bortse från enskilda krav på ett sätt som påminner om stora M-metoden som nämndes i samband med länkekvationerna (se 3.5.3), men är något enklare att förklara. Vi inför helt enkelt en binär variabel, $b_{Trn1,Trn2}$, för varje konflikt mellan två specifika tåg som vi kan tänka oss tillåta. Därefter lägger vi till en term i varje länkekvation som berör de två tågen på följande vis, där K_1 och K_2 är konstanter:

$$\begin{aligned} K_1 b_{Trn1,Trn2} + dep_{Trn1,Lnk} &> dep_{Trn2,Lnk} + tt_{Trn2,Lnk} \\ K_2 b_{Trn1,Trn2} + dep_{Trn2,Lnk} &> dep_{Trn1,Lnk} + tt_{Trn1,Lnk} \end{aligned}$$

Konstanterna K_1 och K_2 väljs medvetet just så att de ska vara tillräckligt stora för att vänsterleden ska bli större än högerleden i ekvationerna ovan om $b_{Trn1,Trn2}$ är 1. Om vi inkluderar $b_{Trn1,Trn2}$ i objektfunktionen på ett ändamålsenligt sätt (se 3.5.8), kommer variabeln att anta värdet 1 endast om det är nödvändigt, dvs om en lösning måste innehålla den specifika konflikten mellan de två berörda tågen. Om den är 1 bidrar förvisso $b_{Trn1,Trn2}$ till kostnaden för lösningen (och kostnaden vill vi som bekant minimera), men effekten blir även att länkekvationerna som upprätthåller avståndet mellan tågen slås ut genom att termerna $K_1 b_{Trn1,Trn2}$ och $K_2 b_{Trn1,Trn2}$ kommer att dominera ekvationerna. Vi har fått en kostnad, men i gengäld kanske det finns en lösning på det återstående problemet, dvs problemet så som det ser ut utan dessa ekvationer.

4 Operatörernas ansökningar

Avregleringen av Sveriges järnväg startade redan år 1988 även om den började i liten skala och i första hand berörde godstågstrafiken, vilken var helt avreglerad i och med juli 1996 [6]. Avregleringen av persontrafiken planeras ske 2010. Idag är samhällsköpt persontrafik konkurrensutsatt.

En av Banverkets uppgifter är att ta in operatörernas ansökningar om hur de vill driva trafik på Sveriges järnvägsnät och se till att kapaciteten på spåren fördelas rättvist. Under nuvarande projekt har vi för första gången fått direkt tillgång till de dokument som tillsammans utgjorde operatörernas ansökningar om trafik på ett visst tidtabellkonstruktionsområde för en viss period (T08). Ansökningshandlingarna kompletterades med korrespondens mellan operatörerna och Banverket via epost och alldeles säkert ett antal telefonsamtal och/eller möten. All sådan korrespondens kan ha modifierat eller förtydligat delar av huvudhandlingarna. Vi har kunnat ta del av åtminstone några dialoger som skett

via epost, medan innehåll i telefonsamtal av naturliga skäl inte gått att åter skapa. I vilket fall som helst har vi haft möjlighet att för första gången kunna studera en färdig tidtabell för ett område och jämföra den med hur operatörerna ursprungligen ville genomföra trafiken.

Även om operatörernas ansökningar inte följer något enhetligt format, uttrycker de i grund och botten önskemål om att få köra en viss mängd tåg, vilket inkluderar att de framställer krav på gångtider eller avgångs- och ankomsttider, eller både och. Vi stöpte mer eller mindre manuellt om operatörernas ansökningshandlingar till ett format som vårt program kan använda. Varje tåg blev därmed förknippat med data som uttrycker när tågen i det ideala fallet ska befinna sig på vilken plats.

En väsentlig del av vad operatörerna egentligen ansöker om kan inte uttryckas med enbart den typen av platskrav som nämns ovan. De har dock inte väsentligt sämre uttryckskraft jämfört med den representation som vi annars använder oss av. T ex finns i ansökningarna implicita krav på gångtider som skillnaden mellan ankomsttid till en plats och avgångstid från den förra platsen. Vi vet dock inget om var smärtgränsen går vad gäller förlängda gångtider relativt de som antyds i ansökningarna, eller i förekommande fall i tidtabellutkasterna vi utgår från. Vi saknar dessutom information om vilka anslutningar operatörerna har tänkt sig. Därför känner vi egentligen inte till om det är acceptabelt eller oacceptabelt för en viss operatör att få ett enskilt tåg flyttat. Detta gäller oavsett om vi utgår från ansökningarna eller från ett tidtabellutkast.

Det finns i dagsläget inget sätt att uttrycka det mindre specifika önskemålet om att få bedriva trafik enligt ett visst mönster, utan ett sådant önskemål uttrycks genom att specificera krav på exakt när tågen ska avgå. Vi kan t ex bara gissa om en grupp tåg i ansökningarna egentligen säger: "vi vill köra tåg en gång varannan timme på det här sättet, men om tågen går 10, 12 och 14 eller 11, 13 och 15 spelar ingen roll".

Största skillnaderna mellan den datamängd som representerar operatörernas ursprungliga ansökningar och kedjorna av traverseringar och platsaktiviteter som vi får från ett tidtabellutkast för samma period och område är att de senare utgör en mer färdig produktionsplan. I princip har tidtabellkonstruktören utgått från samma information som den förstnämnda datamängden representerar för att göra sitt tidtabellutkast, men han eller hon har dessutom haft en dialog med operatörerna och därefter, i samråd med dessa, flyttat på tåg för att komma närmare en körbar produktionsplan.

Vår förhoppning när vi fick tillgång till ansökningarna var att hitta ett sätt att kunna optimera direkt på dessa. Redan de aspekterna som tas upp i föregående stycken antyder att vägen dit är ganska lång. Vi har därför valt att använda ansökningshandlingarna främst som jämförelsematerial för att få svar på specifika frågor om hur den färdiga tidtabellen skiljer sig från de ursprungliga önskemålen – oavsett om tidtabellen är Banverkets egen eller resultatet av optimering från vår sida.

En tänkbar användning är att inkludera vissa aspekter av ansökningarna i optimeringens objektfunktion och på så vis styra slutresultatet mot dessa. Tanken är, bildligt talat, att om vi är fria att gå till höger eller vänster vid ett

visst vägska, och den ursprungliga avsikten var att gå åt höger, är det också vettigare (mindre dyrt) att göra så.

Problemet är hur man ska veta vilka delar av ansökningarna man kan använda på ovannämnda sätt. Det enda som förefaller någorlunda rakt på sak är då ett enskilt tåg har blivit flyttat jämfört med ansökningen och vi vid optimering kan välja mellan att flytta det antingen ännu längre bort från, eller närmare, originalläget. Men inte ens detta fall är oproblematiskt. Hur vet vi att en strävan efter att komma närmare ursprungsläget fortfarande är vettigt om tåg i omgivningen har flyttats? Det kanske finns en anslutning som nu kommer att fungera bättre om vi strävar *bort* från originalläget. Sambanden är minst sagt komplexa och givetvis föremål för pågående forskning [5].

5 Testfall

Syftet med vårt testfall är att i första hand demonstrera det som är nytt i den process som vi föreslår i denna rapport. Med andra ord vill vi med ett konkret exempel visa dels hur åtagandet påverkar produktionsplaneringen, dels hur såväl åtagandet som det resterande planeringsutrymmet kan åskådliggöras i det material som utgör resultatet av produktionsplaneringssteget.

Notera att fokus för testkörningen är produktionsplaneringen. Likväl krävs indata i form av Banverkets åtagande för aktuell period och aktuellt tidtabelläggningsområde för att kunna fullfölja en körning med vårt framtagna prototypverktyg. Vi utgår från att ett åtagande går att ta fram, men beskriver inte här hur det skulle gå till. Det finns givetvis inget historiskt data av denna typ att utgå från eftersom uppdelningen i åtagande och produktion är något nytt som vi föreslår. I brist på autentiskt data har vi därför tagit fram ett material som för vårt ändamål får representera åtagandet.

Komplexiteten hos det problem som vi baserar vårt testfall på har mindre betydelse för demonstrationsvärdet eftersom syftet inte är att demonstrera hur kraftfull vår prototyp är som beräkningsverktyg. Vi har trots detta valt att använda en datamängd vars storlek är jämförbar med ett framtida verkligt problem, dvs mängden tåg motsvarar ungefär förväntat antal tåg i ett tidtabelläggningområde. Däremot har vi i detta test som utgångspunkt valt en konfliktfri tidtabell för tågen i fråga som dessutom innehåller en hel del tidtabellteknisk tid.

Processen kräver två körningar: en för att ta fram och lagra åtagandet i en databas och en för att produktionsplanera med utgångspunkt från åtagandet. Dessa två steg upprepas vid behov om slutresultatet (dvs resultatet från den andra körningen) inte blev tillfredsställande.

5.1 Åtagandet

För att simulera den tidiga processen, har vi låtit en konfliktfri tidtabell⁹ representera operatörernas ansökningar och ur den plockat tidpunkterna för de platsaktiviteter som inbegriper någon form av trafikutbyte, dvs en så kallad trafikaktivitet.

Trafikaktivitet Vi definierar *trafikaktivitet* som en platsaktivitet som uppfyller minst ett av följande kriterier:

- Platsaktiviteten är tågets första eller sista aktivitet
- Platsaktiviteten har ett uppehåll som av TrainPlan klassificeras med någon av följande bokstäver eller bokstavskombinationer: F, FA, FP, G, GA, GP eller X

I ett verkligt fall skulle operatörernas ursprungliga ansökningar naturligtvis inte tillsammans utgöra en konfliktfri tidtabell, men vi vill i detta testfall efterlikna att vi har en fungerande metod för att ta fram ett rimligt åtagande med utgångspunkt från ansökningshandlingarna. Vi beskriver däremot inte metoden i fråga i denna rapport, utan fokuserar på hur krav på avgångar och ankomster bestämmer förutsättningarna för produktionsplaneringen.

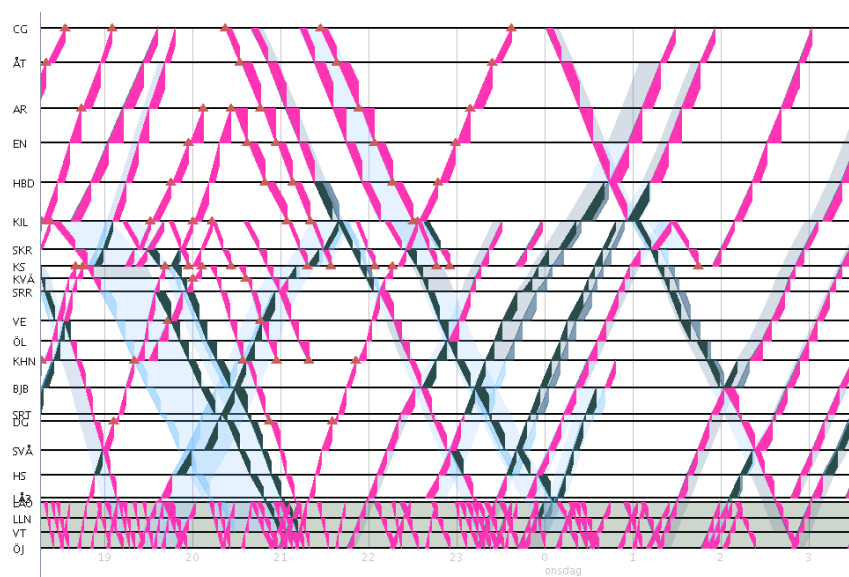
Som nämnt finns det inget egenvärde i att göra problemet svårlöst och vi har därför valt att låta indata vara maximalt tillrättalagt i stället för att komplexitetsmässigt försöka åstadkomma något slags mellanting. Vi är således garanterade minst en konfliktreglerad tidtabell för området i fråga som är förenlig med de tider som gäller för tågens trafikaktiviteter eftersom vi har extraherat nämnda tider ur just en sådan konfliktfri tidtabell.

Triplarna av tidpunkter, tåg och plats för trafikaktiviteterna betraktar vi som Banverkets åtagande och resultatet av den första planeringsfasen i vårt tänkta scenario. Åtagandet är på enklast möjliga form, formulerat så att vi för varje tåg har angivit när det ska ha ankommit till, respektive tidigast får avgå från, platser där det har trafikaktiviteter enligt definitionen ovan.

5.1.1 Körning 1 – Åtagandet

Vi använde den enklaste formen av körning 1 då tiderna för trafikaktiviteterna i det underliggande materialet helt sonika kopieras till databasen. I vårt testfall ingår 1573 tåg och utöver kraven som ställs vid tågens första och sista platser för samtliga tåg finns ytterligare 1053 trafikaktiviteter där ankomst- och avgångstider ska upprätthållas. Det kan påpekas att vissa tåg inte har några sådana trafikaktiviteter alls (i detta tidtabellägningsområde), medan andra tåg är hårt styrda.

⁹I praktiken utgick vi från ett tidtabellkast och strök tåg tills resultatet inte längre innehöll konflikter. Orsaken var att vi inte hade tillgång till en helt konfliktreglerad tidtabell för området i fråga, men behövde ett utgångsmaterial som var fritt från konflikter men samtidigt innehöll en rimlig mängd tidtabellteknisk tid.



Figur 5: Skillnaden mellan ursprungstidtabellen och programmets förslag

Om indata i stället hade varit ett tidtabellutkast som motsvarar det material som en tidtabellläggare har arbetat fram utifrån operatörernas ansökningar, hade vi önskat att körning 1 bearbetade materialet på något sätt för att öka chansen att produktionsplaneringen kan följa åtagandet innan det slutligen läggs i databasen. En första ansats till en sådan metod finns beskriven i arbetspapperet “Påslagsmodellen, Matematisk beskrivning av en grov planeringsmetod för tidtabellkonstruktion” [3].

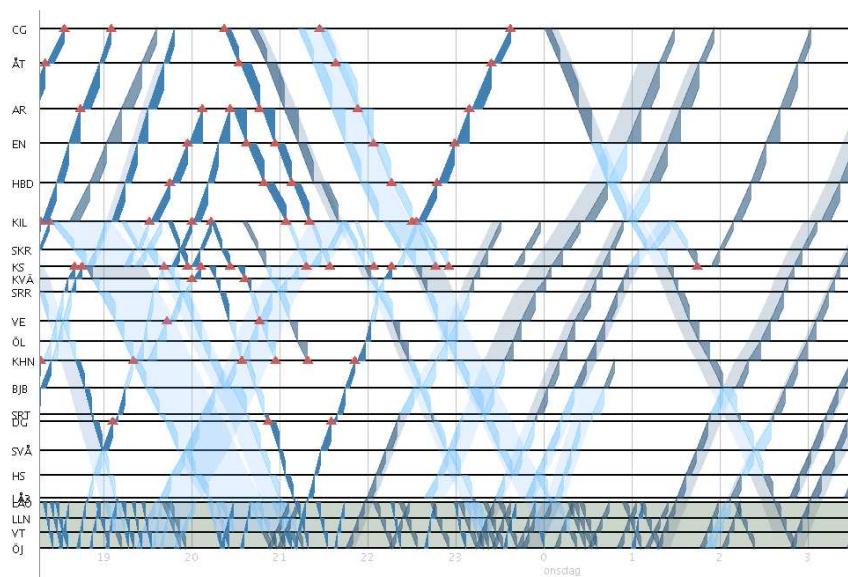
5.2 Produktionsplaneringen

Om man har genomfört körning 1 kan man i körning 2 ange att man vill upprätthålla tiderna för trafikaktiviteterna i den nya körningen så att de stämmer överens med tiderna för motsvarande aktiviteter i databasen, dvs de som genererades under körning 1. Beräkning av intervall sker då enligt den algoritm som beskrivs i 3.3.2, där en omfördelning av tidtabellteknisk tid möjliggörs samtidigt som åtagandet bevaras intakt.

5.2.1 Körning 2 – Produktionsplanering

Underlaget till körning 1 i vårt testfall var en färdig tidtabell och vi väljer i körning 2 att inte slackifiera materialet. Därmed är resultatet från denna körning i princip endast ett tydliggörande av planeringsutrymmet i tidtabellen som användes i körning 1, vilket också var huvudsyftet med testkörningarna.

Pga hur åtagandet är formulerat (“ankommer inte senare än, avgår inte ti-



digare än”) skiljer sig den föreslagna tidtabellen något från den tidtabell som användes som utgångspunkt för körning 1, vilket åskådliggörs i Figur 5: de traverseringar som är identiska för båda tidtabellerna är chockrosa, den ursprungliga tidtabellens traverseringar svarta och den nya tidtabellens traverseringar är blå. Orsaken till skillnaderna är att tidtabellen som programmet föreslår bara är ett av många alternativ till en konfliktfri tidtabell som respekterar åtagandet, och chansen att CPLEX inom de ramar som åtagandet skapar hittar exakt den lösning som den ursprungliga tidtabellen representerar är mycket liten. Men det nya förslaget gör naturligtvis inte heller det några avsteg från åtagandet.

Om vi inte hade hittat en fungerande produktionsplan i detta skede, hade vi i värsta fall kunnat bli tvungna att behöva släppa på delar av åtagandet. För att kunna styra i detalj för vilka par av platser och tåg vi fortfarande ska kräva att åtagandet upprätthålls, behövs explicita villkor för detta. En generell slackifiering upprätthåller som nämnt endast tiderna vid start och slut för tågen (se sid 18).

5.2.2 Om presentationen av resultaten

Emedan det är möjligt, kanske t o m troligt, att man även i fortsättningen vill ha ett detaljerat förslag på hur tågen ska köras, är det viktigt att det plane-ringsutrymme som kvarstår blir tydliggjort för dem som ska tolka resultatet av produktionsplaneringen.

Figur 6 visar en vy av resultatet av vår testkörning. I stora drag är det en vanlig grafisk tidtabell, även om tågens traverseringar ritas som polygoner som

motsvarar deras headway, i stället för som tunna linjer som i dagens grafiska tidtabeller. Helt nytt är dock att det förekommer skuggade områden (*domäner*) runt vissa tåg. Skuggningen visar hur dessa tåg kan köras i stället för på det föreslagna sättet, utan att åtagandet rubbas. Notera även trianglarna (röda) som visar var det finns trafikaktiviteter.

Planeringsutrymmet blir tydligt i en figur av det här slaget. Om en av många tänkbara produktionsplaner väljs och pekas ut som den enda produktionsplanen, utan hänsyn till de (många gånger likvärdiga) alternativ som finns, riskerar planeringsutrymmet däremot att osynliggöras.

Det finns givetvis villkor ställda på tågens aktiviteter som inte framgår av den grafiska representationen som exemplifieras av Figur 6. Det viktigaste (och självklara) villkoret är att ett tåg inte får avgå från en plats innan det har ankommit. Man kan följaktligen inte flytta tågens aktiviteter godtyckligt inom de skuggade områdena. Men vi hävdar att domänerna för tågens avgångar, ankomster och uppehållslängder, tillsammans med en grafisk representation som motsvarar Figur 6, är ett mycket viktigt resultat, och bör följa med det slutgiltiga förslag till tidtabell som levereras till mottagaren av resultatet från produktionsplaneringen.

6 Sammanfattning

Som tidtabellägningsprocessen ser ut idag i Sverige, är tidtabellkonstruktion så komplicerat och tidsödande att arbetet måste utföras innan tillförlitlig information finns att tillgå för att tidtabellkonstruktörerna ska ha en chans att få sina tidtabeller klara i tid. Huvudorsaken är att tidtabelläggning i princip sker manuellt.

Situationen skulle inte vara fullt så allvarlig om processen tillät att tidiga brister i tidtabellen rättades till allteftersom detaljerad och korrekt information blev tillgänglig. Av olika skäl tilläts inte detta i dagens process, vilket naturligtvis resulterar i undermåliga tågtidtabeller som stämmer dåligt överens med verkligheten.

Vi har till Banverket föreslagit en ny process som tar tillvara de möjligheter som datoriserade verktyg kan ge i form av stöd till att mycket snabbare hitta konfliktfria tidtabeller. Vårt förslag till process adresserar en rad faktorer.

Den första viktiga aspekten i det nya processförslaget berör synen på tidtabellen och vad Banverket egentligen ska leverera. Oavsett hur en framtida process ser ut, kommer operatörerna – för att kunna planera sina verksamheter – troligen att behöva få relativt tidiga besked från Banverket om hur de kommer att få köra sina tåg. Dessa besked kommer även i fortsättningen baseras på en planering som sker medan planeringsunderlaget är osäkert och ofullständigt. För att kvaliteten på de slutgiltiga tidtabellerna ska drabbas mindre hårt av detta, föreslår vi att man skiljer på åtagande och produktion. Något förenklat gäller Banverkets åtagande endast en delmängd av avgångarna och ankomsterna i tidtabellen. Hur tågen ska trafikera övriga punkter i nätet blir en produktions-teknisk fråga vars huvudsyfte är att se till att åtagandet uppfylls.

Den andra aspekten gäller Banverkets arbete med att leva upp till sitt åtagande. Produktionsplanerna ska med jämna mellanrum revideras för att inlemma ny kunskap och byta ut felaktiga antaganden mot uppdaterad och mer korrekt information. Fokus ska hela tiden vara hur man bäst ska kunna leverera det man har lovat, dvs hur tågen säkrast ska kunna ankomma respektive avgå vid de tidpunkter som anges i åtagandet. Men även åtagandet ska kunna modifieras om förutsättningarna förändras så att det man ursprungligen lovat blir omöjligt att uppfylla, eller om en operatör så önskar och det finns utrymme att genomföra förändringar utan att åtagandet gentemot övriga operatörer blir lidande. Dock gäller att ingen enskild part ska kunna ändra åtagandet utan att detta förhandlas.

Vidare förespråkar den nya processen att tidtabellkonstruktion bör genomföras med hjälp av optimerande, datoriserade verktyg, även om tidtabellkonstruktören inte helt kan ersättas av en dator i något steg i processen. I den tidiga processen ser vi att de stora dragen fortfarande planeras manuellt som idag eftersom det inte finns något enhetligt och otvetydigt sätt att matematiskt uttrycka alla typer av krav och önskemål på trafiken som operatörerna kan tänkas inkomma med i samband med sina ansökningar. Framförallt finns det inget givet sätt att ställa olika typer av krav mot varandra. Avvägningar av den typen görs bäst av en människa från fall till fall. I det första steget i processen, kan ett datorprogram däremot tänkas behandla de typer av krav som faktiskt går att uttrycka matematiskt och på så vis producera en optimerad plan utifrån dessa, så att tidtabellkonstruktören inte behöver börja med ett blankt blad. I senare skeden av planeringen, lämpar sig datorer utmärkt till att optimera planer inom de allt snävare ramar som tidtabellkonstruktören successivt föreslår, tills resultatet är tillfredsställande.

Sist men inte minst strävar den nya processen efter att hela tiden synliggöra kvarvarande planeringsutrymme. Idag finns det inget sätt att få en överblick över vilka alternativen är under själva tidtabellkonstruktionen, och heller inget sätt att i efterhand se vilka alternativ som stod till buds när tidtabellen fastslogs. Eftersom det heller inte finns någon officiell bild över vilka ankomster och avgångar som utgör åtagandet, är det upp till den enskilda tidtabellkonstruktören eller tågledaren att avgöra från fall till fall vad som är viktigt och mindre viktigt. Den nya processen betonar att resultatet från ett planeringssteg inte bara är en reglerad tidtabell, utan även vilket åtagande man gör och de ramar inom vilka enskilda tåg kan röra sig utan att bryta mot detta åtagande.

Referenser

- [1] Martin Aronsson. Slutrapport för projektet TUFF, TågplaneUtveckling För Framtiden. SICS Technical Report T2006:07, Swedish Institute of Computer Science, 2006.
- [2] Martin Aronsson, Markus Bohlin, and Per Kreuger. Mixed integer-linear formulations of cumulative scheduling constraints - a comparative study. Technical report T2007:03, Swedish Institute of Computer Science, 2007.
- [3] Martin Aronsson and Malin Forsgren. Påslagsmodellen, matematisk beskrivning av en grov planeringsmetod för tidtabellkonstruktion, 2009. Arbetspapper SICS-DDTP 011.
- [4] Banverket. Järnvägsnätsbeskrivning 2010 del 1, bilaga 6 - Tidsplan för tilldelning av kapacitet, December 2008. Nedladdningsbar på http://banportalen.banverket.se/Banportalen/upload/6250/JNB%202010/Bilaga_6_Tidsplan_f%C3%B6r_tilldelning_av_kapacitet_JNB_2010.pdf.
- [5] Per Kreuger, Martin Aronsson, Jan Ekman, and Thomas Franzén. Rail traffic requirements engineering. Technical Report T2005:12, Swedish Institute of Computer Science, 2005. As presented at the 1st International Seminar on Railway Operations, Modelling and Analysis, June, 08-10, 2005, Delft, Netherlands.
- [6] J-E Nilsson. Restructuring Sweden's railways: The unintentional deregulation. *Swedish Economic Policy Review*, 9(2), 2002.